Special Correspondence

A 45 ns Fully Static 16K MOS ROM

JOSEPH WONG, MARK EBEL, AND PHILIP SIU

Abstract—This paper describes a fully static high speed 16 384 bit read only memory (ROM), designed and fabricated by using scaled MOS processing and innovative circuit techniques. Specially designed decoder structures and sense amplifiers enable address accessing in less than 45 ns typically. Extensive use of small signal amplification and 0 volt threshold devices reduce active power to a mere 70 mA, and when the power down option is chosen standby currents reduce to only 15 mA typically. Together with quick turnaround (contact mask programmability), and small die size (20.3K mil^2), this ROM provides a low power, low cost alternative to bipolar PROM's.

I. INTRODUCTION

The high speed read only memory (ROM) market has been traditionally dominated by the bipolar PROM's. However, the majority of them suffer from the disadvantages of excessive power dissipation, as well as high cost due to the inherent complex bipolar technology.

The circuit described here uses scaled metal oxide silicon (MOS) technology to achieve a typical 45 ns address access time performance, very comparable to its bipolar counter parts. However, it only uses 350 MW active power, while the power down option reduces the standby dissipation to a mere 35 MW. Being contact mask programmable to achieve quick turnaround, the die size is still at 130 mil X 156 mil, thus making it a very attractive, low cost replacement for the bipolar PROM's. A side by side comparison of the features of a bipolar PROM and that of this circuit is shown below.

	Bipolar PROM	High speed MOS ROM
Power supply	Single 5 V	Single 5 V
Input and output	TTL compatible	TTL compatible
Typical access time	50 ns	45 ns
Typical active power	650 MW	350 MW
Typical standby power	650 MW	75 MW
Power speed product	32.5 nJ	15.8 nJ
Die size	40 000 mil^2	20 000 mil^2
Number of critical masking steps	10	7

II. CIRCUIT DESIGN

In order to obtain high performance, many circuit tricks are used. All stages play a part in achieving the goal. The following is a stage by stage description of how this is done.

A. Address Buffers

To provide for power down feature, native devices in series with the depletion pull ups are used. The first inverter uses only native pull ups which help to minimize input level sensitivity that can be caused by process variations, such as implant. Since the native transistor threshold is always close to 0 V, an enhancement native first stage will make a much more stable TTL detector than previous enhancement/depletion inverters [1], since enhancement and depletion thresholds vary independently of each other. It is obvious that this will make the ratio of the input buffers much more stable—a very important condition for the 0.8–2.0 V TTL input levels.

Another feature that is built into the address buffer design is that when the power down option is activated, all address lines will remain at about a $V_{cc}/2$ level. This is critical in achieving a fast chip-select accessing time when the circuit is brought back from standby to operating mode, since the outputs are only $\frac{1}{2}$ level from the final value.

A push–pull output stage is used for the address line drivers, together with a small depletion transistor to sustain the steady-state voltage level. This combination will provide the fast rise time needed with the minimum power dissipation—also a very important design goal.

A typical propagation delay through one of these address buffers is 9 ns, while active power dissipation is less than 11 MW; the standby power dissipation, of course, will be equal to the leakage current as a function of the process variations and temperature, but is seldom more than 5–10 μA.

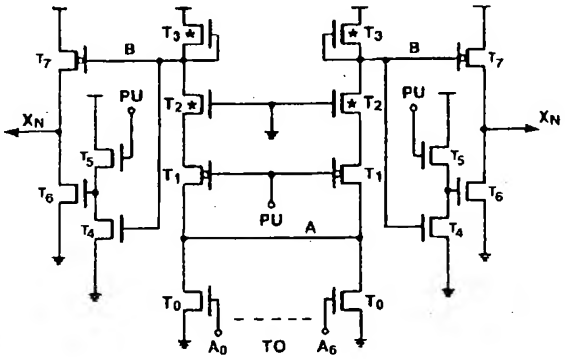


Fig. 1. Decoder driver. PU denotes power up, T_0 is an enhancement device, T_1 is a zero V_T device, and T_2 is a depletion device.

B. Decoder Drivers

This is a very important stage in terms of power–speed product because it possesses the highest number of power dissipating pull ups (127), and has one of the more heavily loaded capacitance nodes—an X-decoder will drive a maximum of 128 transistors. This is a key in power–speed tradeoffs.

A specially designed decoder circuit provides the answer to this challenge. As can be seen in Fig. 1, this circuit uses depletion and native pass transistors as isolation devices between the heavily loaded nor line node A, and the driver node B. Any X-decoder selection allows node A to rise to only one depletion threshold above V_{ss}, while node B swings a full V_{cc}. By using a native push–pull driver, T_3 is not required to deliver large amounts of current. Therefore, decoder current is minimized, while high speed is still maintained. The output stage is powered down high to conserve power up access time.

The typical power dissipation of an X-decoder stage is 1.5 MW, while the propagation delay is approximately 14 ns.

C. Current Sense Amplifier

The sense-amplifier has always been the weak link in MOS ROM's. Since high speed is required, a differential current

Manuscript received April 3, 1981; revised April 23, 1981.
The authors are with Synertek, Santa Clara, CA 95052.
J. Wong is with Synertek, Santa Clara, CA 95052.
M. Ebel and P. Siu are with VLSI Technology, Santa Clara, CA 95052.

0018-9200/81/1000-0592$00.75

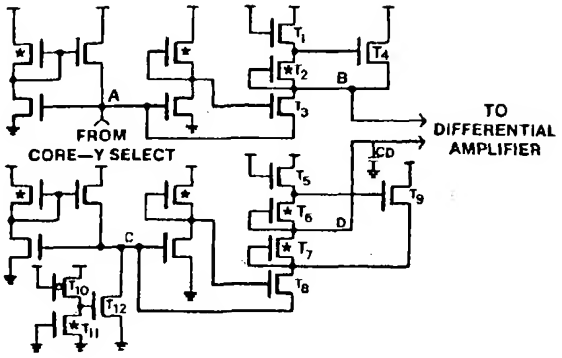


Fig. 2. Sense amplifier.

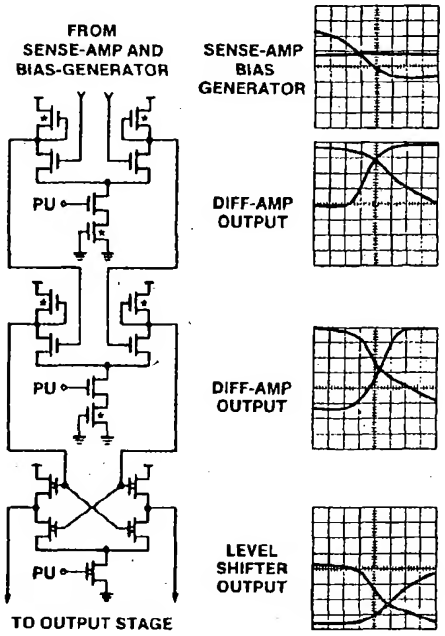


Fig. 3. Differential amplifiers and level shifter.

sensing technique with limited voltage swings, using a dummy cell as a voltage reference, has been developed.

As shown in Fig. 2, the sense-amplifier is basically a current level detection circuit [2] with feedback and clamps to limit the voltage swing of the output to slightly larger than one enhancement threshold. When node A swings approximately 100 MV, node B will swing close to 500 MV, which is the maximum voltage excursion allowed in this design. The operation works in the following manner.

1) Case 1 Node A Conducting. There is a transistor (or bit) in the core for the selected address. The voltage and current level at node A will enable T_3 to turn on, so that the voltage at node B (V_B) is given by

$$V_B = V_{cc} - V_T'(T_1) - V_T'(T_4) - \Delta$$

where

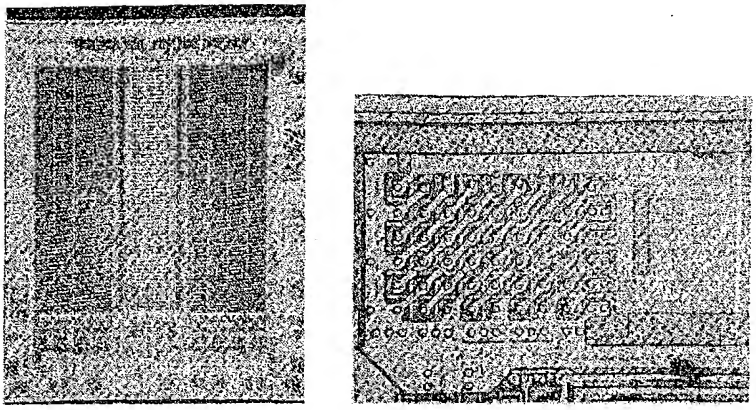
$V_T'(T_1)$ denotes the threshold voltage of enhancement transistor T_3 + body effect due to voltage at node A, $V_T'(T_4)$ denotes the threshold voltage of enhancement transistor T_4 + body effect due to voltage at node B, and Δ denotes the overdrive voltage necessary to sustain the cell current.

2) Case 2 Node A Not Conducting. There is no transistor (or bit) in the core for the selected address. Since there is no current path at node A, voltage bias condition will turn transistor T_3 off, resulting in V_B given by

$$V_B = V_{cc} - V_T'(T_1)$$

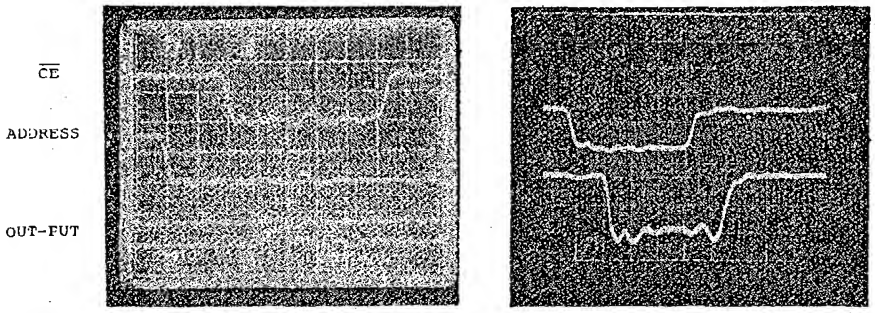
where $V_T'(T_1)$ is similarly defined as above. Thus, it can be shown that node B swings slightly more than one enhancement threshold. Since it does not require large voltage swings, which will add to propagation delay, the advantage of this small signal current sense amplifier is very obvious.

The reference bias generator, also shown in Fig. 2, is a very important element in this design. In order to provide a set of balanced inputs to the subsequent differential amplifiers (see Section II-C), the bias generator is used so that it will generate a reference level at node D, exactly half of that voltage swing at node B. This is done by replacing depletion transistor T_2 in the sense amp with two equivalent transistors T_6 and T_7 in the bias generator, and use node D as an output. The capacitor C_D is added to ensure that there is sufficient capacitance loaded at node D to filter out any cross talk noise.



(a) (b)

Fig. 4. (a) Microphotograph of die. (b) Microphotograph of output device.



(a) (b)

Fig. 5. (a) Chip select access time. (b) Address access time.

It should be pointed out that this bias generator has been designed such that the effect of process variations is minimized. For instance, T_{10} and T_{11} simulate the decoder driver, while T_{12} represents the dummy core bit. All components of this bias generator circuit have been oriented in the same way on the die as their counterparts, so that in almost all aspects it represents an exact replica of the core to sense amplifier path. The propagation delay of this circuit is typically 3–5 ns.

D. Differential Amplifiers

As mentioned in Section II-C, the outputs of the small signal amplifier and bias generator are fed to a series of differential amplifiers [3] (see Fig. 3). The first of these, stage 1, amplifies the 500 MV swing to about 1.0 V, with the high and low levels at about 4.0 and 3.0 V, respectively. These levels are then further amplified, by stage 2, to 4.5 and 2.5 V approximately. Stage 3 is actually a level shifter, the purpose of which is to shift the 2.5 V low level to as close to V_{ss} as possible. The outputs of stage 3 are then fed to the output buffers.

The total propagation delay, as measured from the crossover point of the inputs to stage 1, to that of the output of stage 3, is typically 4 ns.

III. Circuit Performance

A microphotograph of the circuit is shown in Fig. 4. An oscillograph of the typical memory access time is shown in Fig. 5. As mentioned earlier, the typical address and chip select access time is less than 45 ns. The active and standby currents are typically at 70 and 15 MA, respectively. All other electrical parameters are at least as good, if not better, than the bipolar counterparts throughout the temperature and worst case conditions.

IV. Summary

The circuit described here meets the goal of achieving the performance of a bipolar PROM. In addition, it uses less active power, is more economical, and provides a power down feature that is not available in the bipolar circuits. With the flexible contact mask programming capability built into the design, it offers an attractive alternative to the PROM users.

Acknowledgment

The authors would like to thank D. Egashira for the layout, H. Tong for his process support, and L. Lau for his test support.

References

[1] R. Pashley, W. Owen, K. Kokkonen, and A. Ebel, "Speedy RAM runs cooled with power-down circuitry," *Electronics*, vol. 50, pp. 103–107, Aug. 1977.

[2] P. Salsbury, W. Morgan, G. Perlegos, and R. Simko, "High performance MOS EPROMS using a stacked-gate cell," in *ISSCC Dig. Tech. Papers*, 1977, pp. 186–187.

[3] R. Pashley, S. Liu, W. Owen, J. Owen, J. Shappir, R. Smith, and W. Jecmen, "A 16K × 1b static RAM," in *ISSCC Dig. Tech. Papers*, pp. 106–107.

A 20 ns, Low Power, NMOS 1K × 4 Static RAM

CLIFFORD RHODES, RAY PINKHAM, FRED VALENTE, AND RICHARD RAMSEY

Abstract—Using scaled NMOS processing and novel circuit design techniques to enhance the speed-power product, a high-speed, low power, fully static 1024-word × 4-bit random access memory has been

Manuscript received March 6, 1981; revised April 6, 1981.
The authors are with Texas Instruments, Inc., Houston, TX 77001.

0018-9200/81/1000-0594$00.75